

Análise do desempenho do classificador automático MAXVER para uso e cobertura do solo na bacia do rio Mampituba – SC

Micael Barbosa Nadas¹
Tailise Faggion Rodrigues¹
Wladimir Alexandre Pacanaro Trinca¹
Rodrigo Pinheiro Ribas²

¹Universidade de do Estado de Santa Catarina - UDESC
Av. Madre Benvenuta, 2007 - Itacorubi - Florianópolis - SC
CEP: 88.035-901
micael_nadas@hotmail.com

¹Universidade de do Estado de Santa Catarina - UDESC
Av. Madre Benvenuta, 2007 - Itacorubi - Florianópolis - SC
CEP: 88.035-901
tai1287@hotmail.com

¹Universidade de do Estado de Santa Catarina - UDESC
Av. Madre Benvenuta, 2007 - Itacorubi - Florianópolis - SC
CEP: 88.035-901
waptrinca@gmail.com

²Universidade de do Estado de Santa Catarina - UDESC
Av. Madre Benvenuta, 2007 - Itacorubi - Florianópolis - SC
CEP: 88.035-901
rodrigo.ribas@udesc.br

Abstract. This study evaluates the thematic accuracy of the maximum likelihood classifier in a medium spatial resolution imaging satellite Landsat-8. The study area refers to the basin of the Mampituba river in Santa Catarina - Brazil. The analyzed classes were agriculture area, urban area, hydrography, exposed soil and vegetation, where we made a deeper study about the vegetal formations inside the area. The methodology consisted in first of all the discussion about the tools used in image classification such as GIS (Geographic Information System), Remote Sensing and GPS (Global Positioning System) Then, the acquisition of free Landsat 8 images, image processing, classifier training, classification, data analysis and results. The quality of the thematic map was assessed using the kappa statistic, overall accuracy, producer's accuracies and user. The results show that automatic classification given by the classifier gives excellent results for kappa (90,09%) and overall accuracy (93,80%). Among the classes evaluated, the fragment hydrography and bare soil were those with the best accuracies and precisions. The recognition of other classes as agriculture area, urban area, vegetation, depending on the complexity of the landscape and its small dimensions in the study area, depends on the use of image interpretation techniques for further details, making it necessary a new field verification to improve and validate the results.

Palavras-chave: remote sensing, land use, supervised classification, sensoriamento remoto, uso da terra, classificação supervisionada.

1. Introdução

A classificação automática de imagens multiespectrais permite a extração rápida de informação e definição dos padrões de ocupação da terra. Visa à categorização da cobertura da terra distinguindo as composições de diferentes materiais superficiais. Duas abordagens distintas são adotadas na classificação automática de imagens digitais: Classificação supervisionada e não supervisionada. A classificação supervisionada é baseada num conjunto de amostras utilizadas para treinamento do classificador. Portanto, há necessidade de observações de campo na data de passagem do satélite ou conhecimentos de fotointerpretação que permitam o reconhecimento dos padrões de uso da terra. Quando não for possível definir áreas representativas para o treinamento, podem ser utilizados algoritmos baseados em regras estatísticas capazes de fazer o

reconhecimento das classes presentes na imagem. Neste caso a classificação é dita não supervisionada.

Este artigo visa então analisar o desempenho do classificador automático MAXVER, por máxima verossimilhança, com foco em uso e ocupação do solo. A partir do resultado final, será verificada a exatidão do mesmo na classificação dos detalhes contidos nas imagens em relação ao universo real utilizando-se do índice Kappa proposto por Landis & Koch (1977) e índice de exatidão global. Para realizar os procedimentos necessários, utilizaremos o software Spring versão 5.4.2 desenvolvido pelo INPE (Instituto Nacional de Pesquisas Espaciais).

2. Revisão bibliográfica

2.1 Sistema de Informações Geográficas

Segundo LEITE (2013), SIG (Sistema de Informações Geográficas) ou *GIS* (*Geographic Information System*), em inglês, é a ferramenta que permite e facilita a análise, gestão ou representação do espaço e dos fenômenos que nele ocorrem através de *softwares* específicos e que tem como objetivo coletar, armazenar, recuperar, manipular, visualizar e analisar dados espacialmente referenciados a um sistema de coordenadas conhecido.

2.2 Sensoriamento Remoto

De acordo com Jensen (2009), Sensoriamento Remoto consiste no registro da informação de uma determinada região, sem contato, por meio de instrumentos tais como câmeras, escâneres, lasers, dispositivos lineares e/ou matriciais localizados em plataformas tais como aeronaves ou satélites, e a análise da informação adquirida por meio visual ou processamento digital de imagem

2.3 Satélite Landsat 8

Landsat é um programa de satélites de observação da Terra de origem Norte americana. A série teve início na segunda metade da década de 1960, a partir de um projeto desenvolvido pela Administração Nacional de Aeronáutica e Espaço (NASA), sendo dedicado exclusivamente à observação dos recursos naturais terrestres. O satélite Landsat 8 é o oitavo do programa e conta com 7 bandas multiespectrais e uma pancromática, sendo esta com resolução espacial de 15 metros e podendo ser fusionada com as demais tornando assim, todas as imagens com resolução igual à 15m. A obtenção destas imagens é feita de forma gratuita pelo site da USGS (United States Geological Survey).

3. Materiais e Métodos

3.1 Área de Estudo

A área de estudo proposta neste artigo é a bacia hidrográfica do rio Mampituba, uma área de captação de água da chuva que escorre para um rio principal e seus afluentes. Seus limites são definidos pelo relevo, onde os desníveis dos terrenos orientam os cursos d'água e assim determinam a bacia hidrográfica que se forma das áreas mais altas para as mais baixas.

A bacia do rio Mampituba está situada a nordeste do Estado do Rio Grande do Sul entre as coordenadas de 29° 11' a 29° 26' de latitude Sul e 49° 42' a 50° 12' de longitude oeste e no extremo sul do Estado de Santa Catarina (SEMA, 2002), atuando assim como um divisor natural dos dois Estados. O rio nasce nas escarpas basálticas da Serra geral, formação geológica mais importante do Planalto Meridional formada há milhões de

anos devido à grande atividade vulcânica e derramamento de lava. Desemboca no Oceano Atlântico junto à cidade de Torres, após percorrer 62 quilômetros de oeste a leste. Sua vazão média é cerca de 18,6 m³/s com profundidade média de 4 metros e faz a drenagem de uma área de 1.200 km².

O clima regional é do tipo subtropical úmido, com precipitação média anual de 1.502 mm e evapotranspiração média anual apresenta valores de 990 mm (SEMA, 2002). A desembocadura do Rio Mampituba é fixa por molhes, desaguando em uma linha de praia arenosa. Essa porção da zona costeira catarinense encontra-se sob intensa dinâmica de ondas e correntes, e apresenta um regime de micromarés. Está sujeita às ondas mais intensas provenientes de S-SE, que geram um transporte longitudinal resultante ao longo da costa sul catarinense na direção NE (Siegle & Asp, 2007).

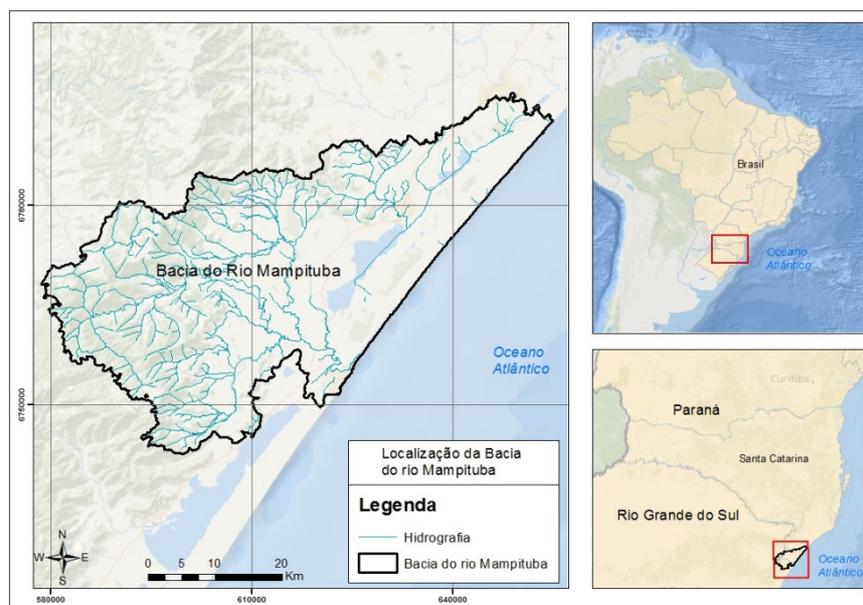


Figura 1. Mapa de localização da bacia do rio Mampituba.

3.2 Metodologia

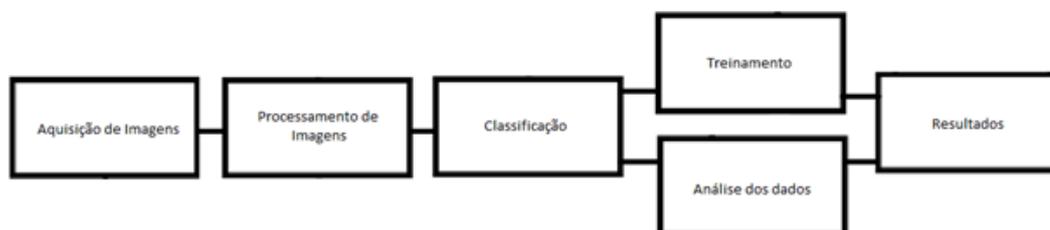


Figura 2. Metodologia adotada na execução do trabalho

Para este trabalho, foram utilizadas as bandas 5, 6 e 4 (RGB) de resolução espacial 30 metros do sensor OLI-TM presente no satélite Landsat 8. As imagens utilizadas são do dia 164 de 2016 da órbita ponto 220 080, adquiridas gratuitamente através do site do Serviço Geológico dos Estados Unidos (USGS). Esta composição se torna interessante

para o mapeamento dentro de uma bacia hidrográfica, pois realça a resposta espectral da vegetação e, se tratando de uma área onde a agricultura é um forte fator socioeconômico juntamente com a predominância de vegetação nativa nas áreas não ocupadas, torna-se mais fácil a identificação de formações vegetais. A vegetação tem uma resposta espectral mais acentuada na faixa do vermelho e infravermelho próximo, tornando assim esta composição a mais indicada para este tipo de mapeamento. Primeiramente, foi feita a restauração das bandas e ajuste de histograma no Spring visando a eliminação de ruídos e informações indesejáveis nas imagens. Posteriormente foi feito o ajuste de filtragem nas imagens, procedimentos estes interessante de serem realizados, pois apresentam uma opção específica para se trabalhar com o sensor OLI-TM do satélite Landsat 8 dentro do Spring, o que gera um resultado altamente confiável. A partir desta etapa, foram geradas composições e imagens sintéticas para posterior escolha de amostras que servem como parâmetro para o algoritmo.

Após todo o tratamento realizado nas imagens descrito anteriormente, foi realizada a classificação supervisionada no software Spring, método este que consiste em primeiramente em criar classes temáticas e adquirir amostras de pixels para reconhecimento do algoritmo em cada classe, rodar de modo automático a análise de amostras com limiar de aceitação em 100%, onde ficará evidente o nível de confusão entre as amostras coletadas para cada classe, podendo assim auxiliar na aquisição de novas amostras se caso haja necessidade. Posteriormente, executar a ferramenta, que irá interpretar a resposta espectral em nível de cinza de cada pixel e as transformará em classes temáticas.

3.3 Classes temáticas

Para este mapeamentos foram utilizadas e adaptadas classes do manual técnico das mudanças na cobertura e uso da terra elaborado pelo IBGE (2015). Neste, as classes apresentam diversos níveis de detalhe diferentes de modo a não se tornar interessante a utilização das mesmas de maneira restrita, mas sim com adaptações.

Água: Inclui-se todas as águas interiores como cursos d'água, canais, sangas, rios corpos d'água naturalmente fechados, lagos, lagoas, lagoas, lagoas e reservatórios artificiais.

Área Artificial: classificada no mapeamento a seguir como área urbana, representa metrópoles, cidades, vilas, áreas de rodovias, serviços e transportes, redes de energia, comunicações e terrenos associados, áreas ocupadas por indústrias, complexos industriais e comerciais e edificações que podem, em alguns casos, estar situadas em áreas peri-urbanas.

Área Agrícola: nomeada como cultura no mapeamento a seguir, pode ser definida como área utilizada para produção de alimentos, fibras e commodities do agronegócio. Inclui terras cultivadas, caracterizadas pelo delineamento de áreas cultivadas ou em descanso, podendo também compreender áreas alagadas cultivadas.

Vegetação Florestal: nomeada como simplesmente vegetação neste mapeamento, caracteriza-se por formações arbóreas, com porte superior a 5 metros de altura, incluindo-se ai as áreas de Floresta Densa, Floresta Aberta, Floresta Estacional, além da Floresta Ombrófila Densa de Terras Baixas, Submontana, Montana e Alto Montana, Floresta Nebular, Floresta Ombrófila Mista e Estepes (Campos). Situada entre o planalto e o oceano, a Floresta Ombrófila Densa é constituída por um estrato arbóreo que pode chegar até 30 m de altura, com copas largas e folhagem perenefoliada, como a Canela Preta (*Ocotea catharinensis*), já em seu estrato mediano encontra-se árvores de porte médio, como o palmitero (*Euterpe edulis*), também possui um estrato arbustivo com predominância de espécies dos da família das *Rubiáceas*, *Palmáceas* e *Monimiáceas*, além de uma enorme quantidade Bromélias e Orquídeas. (SANTA

CATARINA, 1986). A Floresta Ombrófila Densa de Terras Baixas abrange “as florestas de planícies quaternárias costeira de origem fluvial e flúvio marinha [...] situadas em altitudes desde o nível do mar até aproximadamente 30 m” (SANTA CATARINA, 1986), é uma floresta pouco desenvolvida e pouco densa. Já a Sub-Montana e a Montana situam-se entre 400 e 1000 m de altitude, no entanto a Sub-Montana ocupa mais a área litorânea e a Montana ao longo da Serra do Mar e da Serra Geral, ambas com predominância da *Ocotea catharinensis* (Canela Preta). A Alto Montana está localizada em níveis acima de 1000 m de altitude, ocupando as encostas das Serras do Mar e Geral e forma uma vegetação mais aberta, com uma composição florística mais distinta. Segundo Santa Catarina (1986), a Floresta Ombrófila Mista é uma “floresta particularmente restrita ao planalto”, ocorre em altitudes acima de 500 m. Em seu estrato arbóreo, predomina a *Araucária angustifolia* (pinheiro do paraná) e no estrato das arvoretas, predomina a *Illex paraguariensis* (erva-mate). Essa floresta também é conhecida como Floresta de Araucárias. Ainda segundo Santa Catarina (1986), a Estepe (Savana ou Campos) é formada por um estrato de gramíneas de variadas espécies, dentre elas o capim-caninha (*Andropogon lateralis*) e a grama-forquilha (*Paspalum notatum*), ocorre em altitudes acima de 800 m, em terrenos lixiviados ou aplainados, dependendo da constituição geológica. Esta formação vegetal coexiste com a Floresta Ombrófila Mista.

Áreas Descobertas: denominada de solo exposto neste trabalho, compreende tanto áreas de pastagem como também afloramentos rochosos, recifes, áreas com processo de erosão ativo, formação de dunas litorâneas e interiores, dunas móveis sem vegetação, dunas estabilizadas e campos de dunas continentais.

4. Resultados e Discussões

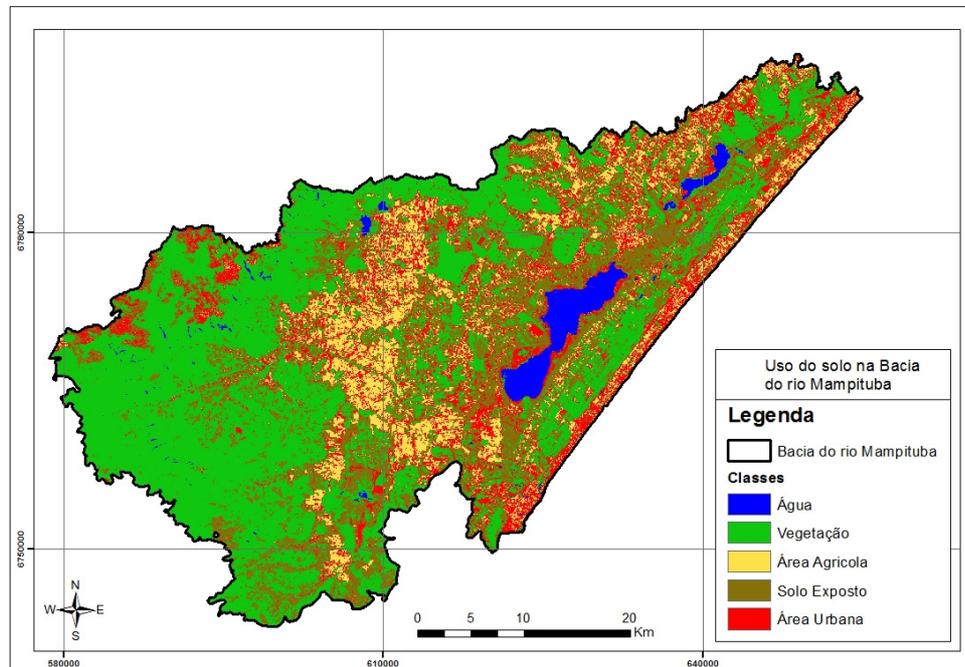


Figura 3. Resultados da classificação automática utilizando o algoritmo MAXVER.

O classificador utilizado, MAXVER, considera a ponderação das distâncias entre médias dos níveis digitais das classes, utilizando parâmetros estatísticos. Para que a classificação por máxima verossimilhança seja precisa o suficiente, é necessário um número razoavelmente elevado de "pixels", para cada conjunto de treinamento. Os

conjuntos de treinamento definem o diagrama de dispersão das classes e suas distribuições de probabilidade, considerando a distribuição de probabilidade normal para cada classe do treinamento.

Após o resultado, o software gera um arquivo chamado Matriz de Erros de Classificação, onde são apresentados os dados referentes as amostras analisadas quanto a nível de confusão entre as classes e porcentagem de qualidade de resultado utilizando-se do índice Kappa.

Tabela 1. Resultados do mapeamento em relação à exatidão obtida.

	Exatidão do produtor	Exatidão do usuário
Água	100.00%	99.93%
Vegetação	99.70%	94.55%
Cultura	64.98%	85.42%
Solo_Expos	39.97%	96.59%
Área_Urban	83.44%	60.77%

Segundo Moreira (2007), a qualidade de um mapa temático é avaliada com base em dois quesitos: a precisão e exatidão ou acurácia de mapeamento. Enquanto que a precisão do mapeamento significa quanto à área de cada classe reflete a verdade no campo, a exatidão é um número (índice) que avalia o posicionamento da distribuição espacial de cada uma das classes que foram mapeadas. Segundo Congalton (1991) uma das técnicas mais utilizadas na realização de avaliação de acurácia da classificação de dados de sensoriamento remoto é o emprego da matriz de erro, também conhecida como matriz de confusão.

Na tabela 1, fica evidente onde mais houve confusão do classificador ao analisar as amostras fornecidas pelo usuário, onde as classes de cultura e área urbana tiveram os menores percentuais de exatidão, logo, maior nível de confusão. Esta confusão se dá principalmente pela semelhança da resposta espectral entre os alvos das duas áreas, onde o classificador apresenta dificuldades para diferenciar as duas.

```

Desempenho geral: 93.80 %
Confusao media : 6.20 %
Abstencao media : 0.00 %
Estistica KHAT : 90.09 %
Variancia KHAT : 2.870e-006
Estistica TAU : 92.25 %
    
```

Figura 4. Desempenho geral e índices avaliativos da classificação.

O índice Kappa, representado na tabela 2 como Estatística KHAT, resultou em valor de 90,09%, e exatidão global de 93,80%, representada por Desempenho Geral, sendo uma avaliação excelente de acordo com a tabela de referência de Lands & Koch,

(1977). Segundo Lillesand (2004) o índice Kappa serve como um indicador do grau que a porcentagem dos valores corretos da matriz de erro é devido à “verdade” versus a “probabilidade” de concordância. Neste caso, o índice Kappa obtido de 90% deve ser entendido como uma indicação de que a classificação obtida foi 90% melhor do que seria um resultado originado através de uma assinatura aleatória dos pixels.

Tabela 2. Índice de qualidade Kappa. Percebe-se que o resultado deste mapeamento classifica-se como Excelente.

Valor de Kappa	Qualidade do mapa temático
< 0,00	Péssima
0,00 – 0,20	Ruim
0,20 – 0,40	Razoável
0,40 – 0,60	Boa
0,60 – 0,80	Muito Boa
0,80 - 100	Excelente

Adaptada de Landis e Koch (1977).

5. Conclusões

Os resultados mostram que a classificação automática conferida pelo classificador MAXVER apresenta excelentes resultados de índice Kappa (90,09%) e de exatidão global (93,80%). Dentre as classes avaliadas, as de fragmento água e solo exposto foram aquelas que apresentaram as melhores exatidões e precisões. O reconhecimento das demais classes como área agrícola, área urbana e vegetação em função da complexidade da paisagem e de suas pequenas dimensões na área de estudo, depende do uso de técnicas de fotointerpretação e visita pontual a campo para maiores precisões no resultado final do mapeamento.

Ressalta-se que a partir de uma breve pós-classificação é possível obter um resultado mais satisfatório, de modo que este procedimento irá generalizar e incorporar polígonos muito pequenos à polígonos maiores. Esta pós-classificação pode ser realizada tanto manualmente como automaticamente dentro do software Spring.

REFERÊNCIAS

COLGATON, R. G. **A review of assessing the accuracy of classifications of remotely sensed data.** *Remote Sensing of Environment*, v. 49, n. 12, p. 1671-1678, 1991.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: uma perspectiva em Recursos Terrestres.** Tradução de J. C. N. Epiphânio. São José dos Campos, SP: Parênteses, 2009. 598 p. (Prentice Hall Series in Geographic Information Science) Tradução de: Remote Sensing of the environment: na earth resource perspective.

Landis, J.R.; Koch, G.G. **The measurement of observer agreement for categorial data.** *Biometrics*, v. 33, p. 159- 154, 1977

LEITE, Marcos Esdras (org). **Geotecnologias aplicadas aos estudos geográficos.** Montes Claros: Unimontes, 2013.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W. **Remote sensing and interpretation.** 5ed. Madison: Wiley, 2004. 763p.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação.** 3.ed. Viçosa, MG: UFV, 2007. 320p.

SANTA CATARINA. Gabinete de Planejamento e Coordenação Geral. Subchefia de estatística, geografia e Informática. **Atlas de Santa Catarina.** Rio de Janeiro, Aerofoto Cruzeiro, 1986

SIEGLE E & ASP NE. 2007. **Wave refraction and longshore transport patterns along the southern Santa Catarina coast.** *Braz. J. Oceanogr.*, 55(2): 109-120.